Manufacturing Letters 37 (2023) 53–56

Contents lists available at ScienceDirect

Manufacturing Letters

journal homepage: www.elsevier.com/locate/mfglet

Letters

Accelerated and inexpensive Machine Learning for manufacturing processes with incomplete mechanistic knowledge

Jeremy Cleeman, Kian Agrawala, Rajiv Malhotra ⇑

Department of Mechanical and Aerospace Engineering, Rutgers University, 98 Brett Road, Piscataway, NJ, USA

a r t i c l e i n f o

Article history:
Received 9 April 2023
Received in revised form 5 July 2023
Accepted 10 July 2023
Available online 6 August 2023

Keywords:
Machine Learning
Transfer learning
Data generation cost
Manufacturing processes
Fused Filament Fabrication

a b s t r a c t

Machine Learning (ML) enables deployable modeling of parametric effects in manufacturing processes. But this paradigm is largely limited to established processes, since the state-of-the-art ignores the significant cost of creating qualitatively accurate physics-based models for new processes. We propose a transfer learning based method that addresses this issue by pushing the boundaries of the qualitative accuracy demanded of the physics-based model. Our approach is evaluated for modeling the printed line width in Fused Filament Fabrication and shows reduction in the model development cost by multiple human-years, experimental cost by 56–76%, computational cost by orders of magnitude, and error by 16–24%.

1. Introduction

Machine Learning (ML) has become popular for modeling parametric effects in manufacturing processes due to its high deployability. But generating the required training data from experiments incurs time and resources (experimental cost C_E). Generating the training data from physics-based process models incurs a computational cost C_C, i.e., the CPU-hours needed to run simulations; and a model development cost C_D, i.e., the time and human resources needed for intuitive trial-and-error creation of constitutive laws and numerical methods that qualitatively and quantitatively capture interactions between multiple physical phenomena over multiple time and length scales [1]. The root cause of high C_D for new processes, which is commonly on the order of decades [2–5], is that qualitative knowledge of the underlying physics is often missing.

Multifidelity learning trains an ML model using a large amount of inexpensive and inaccurate data (source) and fine-tunes it using a small amount of costly but accurate data (target). Using computational process models as the source and experimental data as the target reduces C_E relative to training with only experimental data, reduces C_C compared to training with only computational data, and captures the ground truth [6]. But this approach assumes that the source must qualitatively match the target, i.e., multifidelity learning only effects a quantitative correction. Thus, C_D is still high since a qualitatively accurate physics-based source is needed. Using analytical process models as the source decreases C_C even further, but does not reduce C_D [7]. Note that using experimental sources for new processes is not possible due to their inherent novelty.

This paper proposes a multifidelity learning approach that reduces C_D despite limited mechanistic knowledge of the process physics. This method is demonstrated for modeling the printed line's width W in Fused Filament Fabrication (FFF) as a function of the filament feed rate F and extruder speed S. This problem involves complex physics including non-Newtonian flow, friction, cooling, wetting, and compressibility [8,9]. While FFF is not new and this problem has an existing solution, we choose this problem since this very fact that allows us to quantify the reduction in C_D possible.

2. Methods

Our approach uses transfer-based multifidelity learning, with a physics-based process model as the source and experiments as the target. Thus, the final ML model reflects the experimental ground truth and reduces C_E. This process model must (a) include one or more conservation laws to respect the fundamental laws of nature; (b) use a guess for the form of the constitutive law without any experimental calibration or validation, to reduce C_D; (c) avoid or minimize spatiotemporal discretization to minimize C_C. The reader is referred to the literature for the various transfer learning methods available for regression [10,11]. In this paper, Epsilon Support Vector Regression (SVR [12,13]) with a gaussian Radial Basis Function was used as the ML model and the TrAdaBoost.R2

⇑ Corresponding author.
E-mail address: rajiv.malhotra@rutgers.edu (R. Malhotra).

https://doi.org/10.1016/j.mfglet.2023.07.017

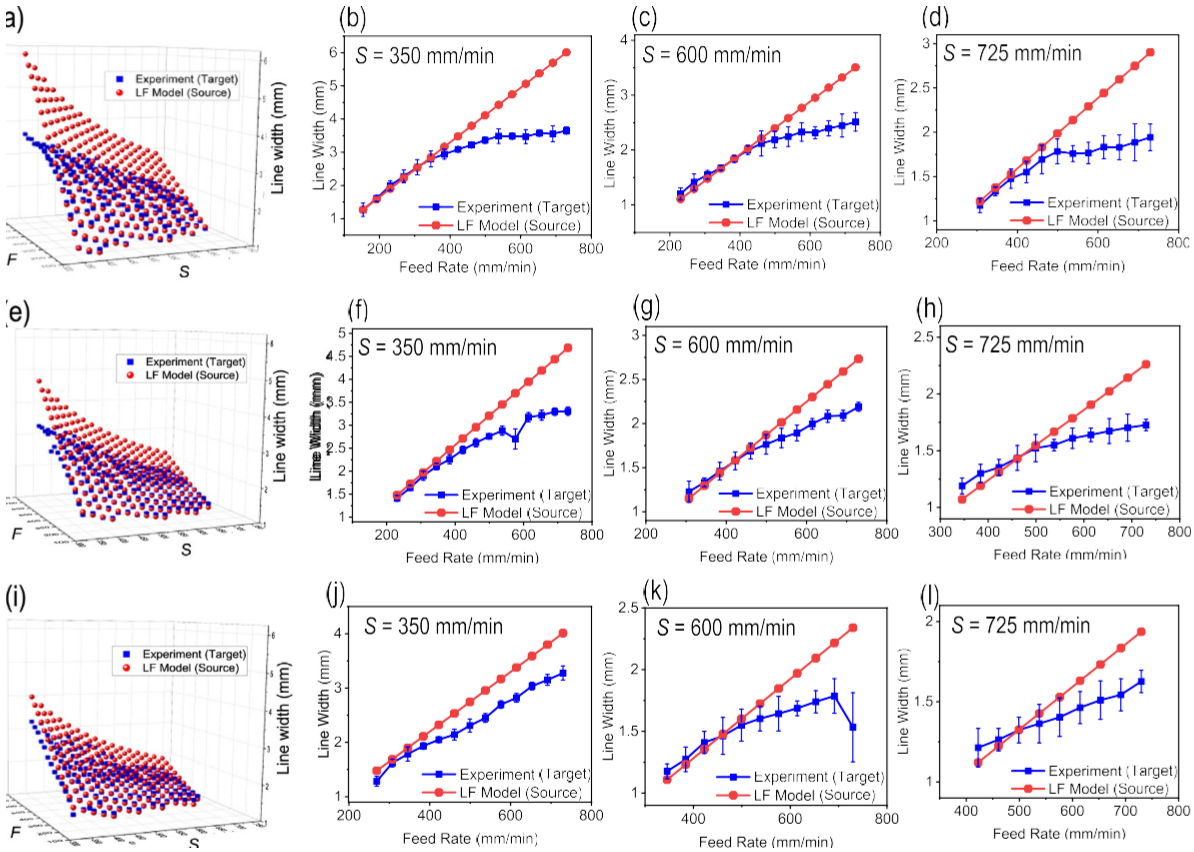

Fig. 1. Comparison of source and target for h = (a-d) 0.7 mm (e-h) 0.85 mm (i-l) 1.2 mm. Feed rate F and stage speed S are in mm/min.

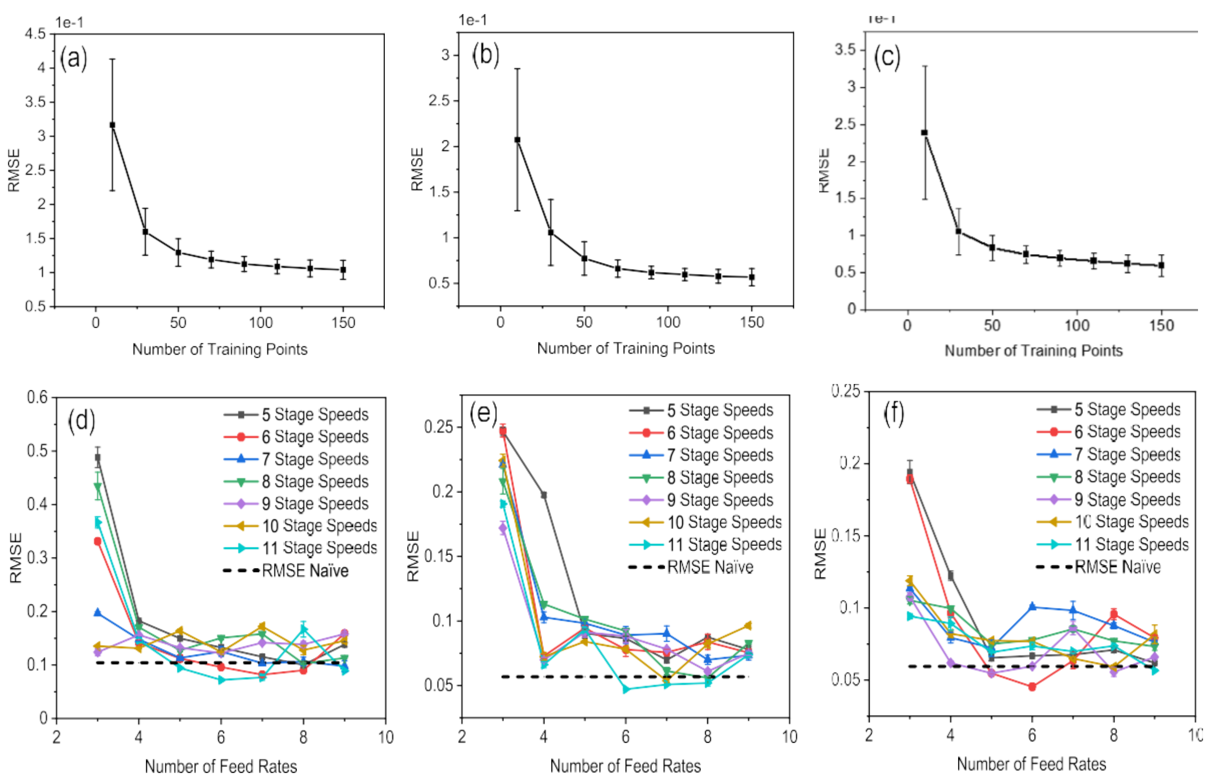

Fig. 2. RMSE from naive learning as a function of the number of target training points for h = (a) 0.7 mm (b) 0.85 mm (c) 1.2 mm. Comparison of $RMSE_{naive}$ to the error obtained from transfer learning using different amounts of experimental F and S and for h = (d) 0.7 mm (e) 0.85 mm (f) 1.2 mm. Feed rate F and stage speed S are in mm/min.

algorithm was used for transfer learning [14]. The hyperparameters for the e-SVR were based on brute force identification and the number of boosting iterations for TrAdaBoostR2 was 30. The reader is referred to the above literature for further details on both SVRs and TrAdaBoostR2.

The source model was the mass conservation law, i.e., $W = FA/Sh$, where h is the nozzle-to-platen distance and A is the filament's cross-sectional area. This model ignores the complexity of almost all the earlier mentioned extrusion physics, and makes an incorrect but simplifying assumption that h equals the line (or layer) height. It took � 10^{-6} CPU-hours to generate the 624 source samples used here. Experiments were performed to print PLA lines on a home-built FFF machine with a 1 mm diameter nozzle for sixteen equidistant S (between 350 and 725 mm/min) and F (between 153 and 729 mm/min) across h = 0.7, 0.85, 1.2 mm. The W was measured using vernier calipers and averaged across 3 measurements. Unstable printing regimes were excluded.

First, naive learning of the SVR was performed on only the experimental target data. Progressively more training points were used till the Root Mean Square Error (RMSE) on the testing data (i.e., the remainder of the dataset) did not decrease further. This training and testing was performed 1000 times using random sampling, and yielded the average values of the smallest error $RMSE_{naive}$ and the corresponding number of samples n_{naive}. Transfer learning was performed with the source data of the same size as n_{naive}. A progressively increasing amount of target data was used to iteratively identify the smallest target dataset needed for transfer learning (n_t) such that the transfer learning error $RMSE_t$ was lesser than or equal to $RMSE_{naive}$. This ensured that prediction accuracy was not sacrificed in the drive to reduce C_D. Testing of the final SVR obtained after transfer learning was performed on data obtained randomly from the a-priori withheld portion of the experimental dataset. This test dataset was of the same size as n_{naive} to prevent a heavily lopsided train:test ratio and thus fairly compare naive and transfer learning. This randomized testing was performed 30 times to obtain the mean $RMSE_t$.

3. Results

Fig. 1 shows the functional discrepancy between the source model and the experimental target with 3D plots and representative 2D plots. The true effect of F and S on W is decidedly nonlinear, especially at lower h, as compared to the linear assumption in the source. Fig. 2a-c show the change in the tested RMSE of naive learning on only experimental data as a function of the number of training points, and reveals the $RMSE_{naive}$ and n_{naive} (which is constant at 150 for all h). Fig. 2d-f compare this $RMSE_{naive}$ to the error from transfer learning for different amounts of experimental training data (i.e., combinations of F and S). There are multiple combinations of F and S for which transfer learning enables $RMSE_t$ ::; $RMSE_{naive}$ and $n_t < n_{naive}$. Qualitatively, Fig. 3 shows that the transfer learnt SVR can capture the nonlinearity in the experimental data despite the qualitatively and quantitatively inaccurate mechanistic knowledge embedded in the source model.

Our approach realizes a 56–76 % reduction in C_{EXP} as compared to naive learning and reduces the error by 16–24% (Table 1). Existing computational or analytical process models can be used as the source for naive learning since they are good qualitative

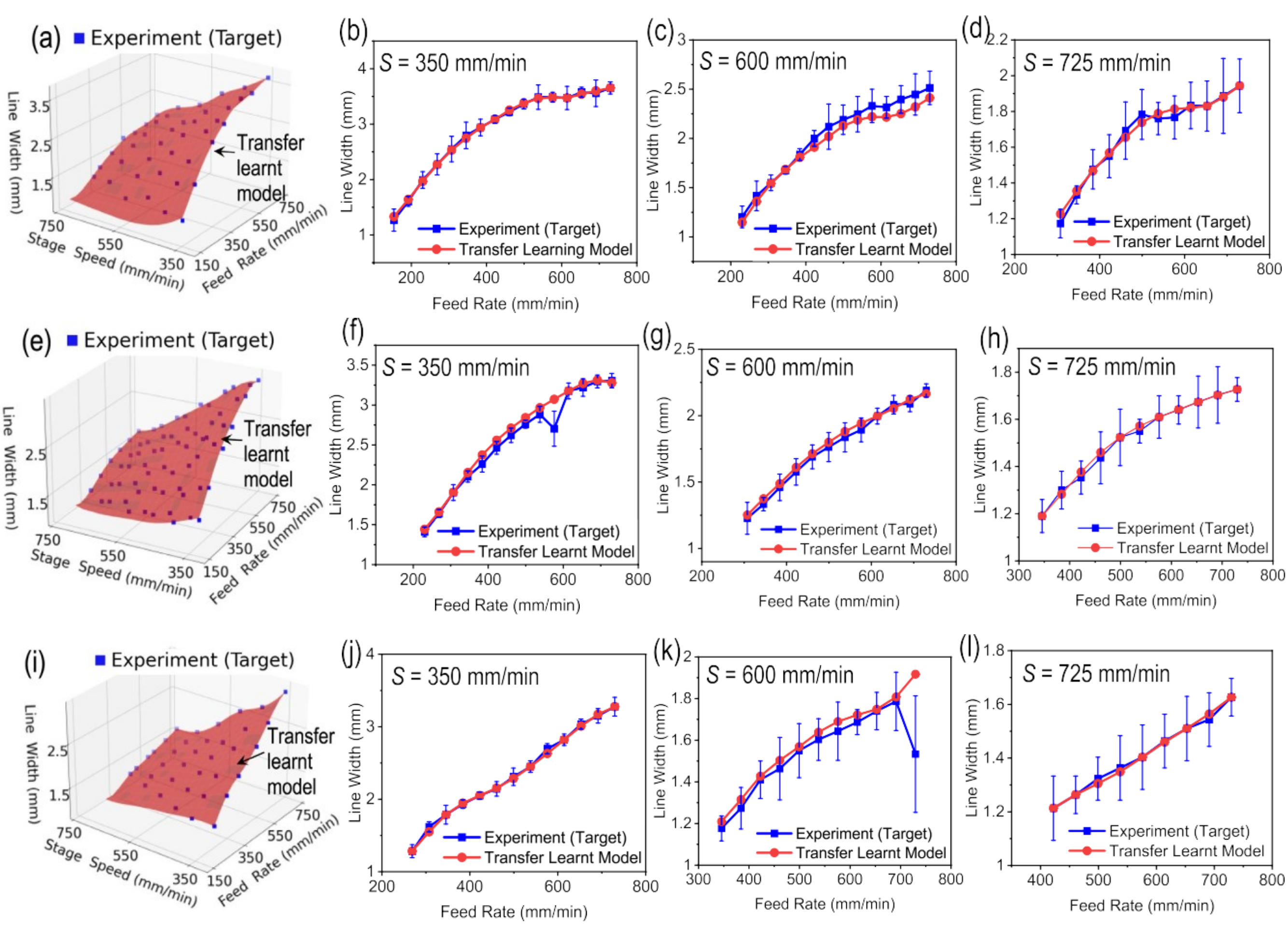

Fig. 3. Comparison of transfer learnt model and target for h = (a-d) 0.7 mm (e-h) 0.85 mm (i-l) 1.2 mm.

Table 1
Comparison of smallest RMSE and corresponding number of training samples for naive learning and transfer learning.

h (mm)	n_{naive}	$RMSE_{naive}$	n_t		$RMSE_t$	$\frac{n_{naive}-n_t}{n_{naive}}$	$\frac{RMSE_{naive}-RMSE_t}{RMSE_{naive}}$
			No. of S	No. of F			
0.7	150	0.104 ± 0.014	6	7	0.081 ± 0.004	72%	22%
0.85	150	0.056 ± 0.009	11	6	0.047 ± 0.0006	56%	16%
1.2	150	0.059 ± 0.015	6	6	0.045 ± 0.002	76%	24%

and quantitative matches to the ground truth [8,9]. But it has taken significant time and effort for these models to reach this point, from 2000 to 2019 for analytical equations [15,16] and from 2002 to 2018 for computational simulations [9,17]. This indicates that using Smart-ML in 2000, which is when our source model was reported in the literature, could have saved at least 18 human-years of C_{DEV}. Overall, our approach reduces C_{DEV} for new processes by easing the need for qualitatively accurate human-created physics-based process models. Note that using high-fidelity computational models to generate just one training sample for FFF needs orders of magnitude more CPU-hours than that for Smart-ML (i.e., 10^{-6} CPU-hours) [9,16]. Thus, Smart-ML reduces C_{DEV} in addition to C_{COMP} and C_{EXP}.

4. Conclusions

State-of-the-art approaches for ML models of parametric effects in manufacturing processes focus on reducing the experimental and computational cost of training data generation. This paper pushes beyond this paradigm to examine the possibility of also reducing the often-overlooked, but significant, cost of process model development. This is achieved by testing the limits of the requisite similarity between source process models and target experimental data in transfer learning, via the use of an uncalibrated guess for the functional form of the constitutive law to avoid the cost of iterative model development. This approach overcomes significant functional discrepancies between the source and the target, unlike assumptions made in the manufacturing literature; reduces the developmental cost along with the experimental and computational costs of generating training data; and reduces the prediction error. A key challenge that will be tackled in our future work is automated and intelligent sampling of the experimental target, since the number of process parameters can be numerous and where to sample in this parameter space is not intuitive for a new manufacturing process.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

All the authors acknowledge the support from the US National Science Foundation Award No. CMMI # 2001081. Jeremy Cleeman further acknowledges additional support from the US National Science Foundation Graduate Research Fellowship Program (Fellow ID: 2022291313)

References

[1] Arinez JF, Chang Q, Gao RX, Xu C, Zhang J. Artificial intelligence in advanced manufacturing: current status and future outlook. J Manuf Sci Eng 2020;142.
[2] Allison J, Li M, Wolverton C, Su XuMing. Virtual aluminum castings: an industrial application of ICME. JOM 2006;58(11):28–35.
[3] Mukherjee T, DebRoy T. A digital twin for rapid qualification of 3D printed metallic components. Appl Mater Today 2019;14:59–65.
[4] Zoch H-W. From single production step to entire process chain – the global approach of distortion engineering. Mater Werkst 2006;37(1):6–10.
[5] Zoch H-W. Distortion engineering – interim results after one decade research within the Collaborative Research Center. Mater Werkst 2012;43(1-2):9–15.
[6] Alam MF, Shtein M, Barton K, Hoelzle DJ. Autonomous manufacturing using machine learning: a computational case study with a limited manufacturing budget V002T007A009pages (American Society of Mechanical Engineers).
[7] Liu S, Lu Y, Zheng P, Shen H, Bao J. Adaptive reconstruction of digital twins for machining systems: a transfer learning approach. Rob Comput Integr Manuf 2022;78:102390.
[8] Serdeczny MP, Comminal R, Pedersen DB, Spangenberg J. Experimental and analytical study of the polymer melt flow through the hot-end in material extrusion additive manufacturing. Addit Manuf 2020;32:100997.
[9] Serdeczny MP, Comminal R, Pedersen DB, Spangenberg J. Experimental validation of a numerical model for the strand shape in material extrusion additive manufacturing. Addit Manuf 2018;24:145–53.
[10] Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. A comprehensive survey on transfer learning. Proc IEEE 2021;109(1):43–76.
[11] Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. J of Big data 2016;3:1–40.
[12] Jain RK, Smith KM, Culligan PJ, Taylor JE. Forecasting energy consumption of multi-family residential buildings using support vector regression: investigating the impact of temporal and spatial monitoring granularity on performance accuracy. Appl Energy 2014;123:168–78.
[13] Drucker H, Burges CJ, Kaufman L, Smola A, Vapnik V. Support vector regression machines. Adv Neural Inf Proces Syst 1996;9.
[14] Pardoe D, Stone P. In: Proceedings of the Twenty-Seventh International Conference on Machine Learning, ICML 10 (Haifa, Israel, 2010).
[15] Bellini A, Gu¨c_eri S, Bertoldi M. Liquefier dynamics in fused deposition. J Manuf Sci Eng 2004;126(2):237–46.
[16] Agassant J-F, Pigeonneau F, Sardo L, Vincent M. Flow analysis of the polymer spreading during extrusion additive manufacturing. Addit Manuf 2019;29:100794.
[17] Bellini A. Fused deposition of ceramics: a comprehensive experimental, analytical and computational study of material behavior, fabrication process and equipment design. Philadelphia, PA: Drexel University; 2002. PhD dissertation.